The Effect of Resampling on the Trace-Off of Resolution and Variance
Bill Menke, 5/17/19

These ideas are based on a discussion that Dan Blatter and I had on the effect of resampling on the trade-off of resolution and variance, as computed by the method we recently developed. The discussion here imagines that a coarsely-sampled time series \( m \) is the fundamental one (e.g. computed via an MCMC method) but that a finely-sampled time series \( y \), computed from \( m \) by interpolation, is the one upon which the resolution/variance analysis is based. The question is how much bias is introduced by having done the resolution/variance analysis in \( y \) rather than \( m \).

I analyze this question below and provide the answer “not much”.

My identification of \( m \) as a coarsely-sampled time series representing some underlying continuous function \( m(x) \) is a simplification. It is meant to stand-in for the more interesting – and more complicated - case, not treated here, in which \( m \) is the set of coefficients of a spline representation of \( m(x) \).

1. Eigenvalues and Eigenvectors of a “Modulated Repeating Pattern” Matrix. The \( N \times N \) matrix \( B \) is constructed by repeating a \( P \times P \) block \( W \) on an \( M \times M \) grid, so that that \( N = PM \). Each instance of \( W \) is multiplied by a factor given by the elements of an \( M \times M \) symmetric matrix \( A \), so that:

\[
B_{ij} = A_{(i/P),(j/P)} W_{(i/P),(j/P)}
\]

(1)

Here matrix indices start at zero, “/” is integer division and “\( \)\" is remainder. We assume that \( A \) and \( W \) satisfy the algebraic eigenvalue equations, \( Aa(s) = \alpha_k a(s) \) and \( Ww(s) = \omega_k w(s) \), respectively. We can show that the quantities:

\[
\beta_s = \alpha_{(s/P)}\omega_{(s\backslash P)} \quad \text{and} \quad b_{k}^{(s)} = a_{(k/P)}^{(s/P)} w_{(k\backslash P)}^{(s\backslash P)}
\]

(2)

are the eigenvalues and eigenvector of \( B \) by considering the product \( Bb^{(k)} \):

\[
[Bb^{(s)}]_q = \sum_{k=0}^{N-1} B_{q,k} b_{k}^{(s)} = \sum_{k=0}^{N-1} A_{(q/P),(k/P)} W_{(q\backslash P),(k\backslash P)} a_{(k/P)}^{(s/P)} w_{(k\backslash P)}^{(s\backslash P)}
\]

(3)

With the substitution, \( k = iP + j \), the single sum \( k = 1 \cdots N \) can be replaced by two sums, \( i = 1 \cdots M \) and \( j = 1 \cdots P \). Then, noting that \( k/P = i \) and \( k\backslash P = j \), we have:

\[
[Bb^{(s)}]_q = \sum_{i=0}^{M-1} \sum_{j=0}^{P-1} A_{(q/P),i} W_{(q\backslash P),j} a_{i}^{(s/P)} w_{j}^{(s\backslash P)} =
\]
\[
\sum_{i=0}^{M-1} A_{q/p,i}a_i^{(s/p)} \sum_{j=0}^{P-1} W_{q\setminus p,j}w_j^{(s\setminus p)} = \\
= \alpha_{(s/p)}\alpha_{(s\setminus p)} \sum_{i=0}^{P-1} a_i^{(s/p)}w_i^{(s\setminus p)} = \beta_k b_k^{(s)}
\]

(4)

Thus, the choices above satisfy the algebraic eigenvalue equation \(BBb^{(s)} = \beta_k b^{(s)}\).

2. Eigenvalues and Eigenvectors of a “Block-Constant” Matrix. Now consider the special case where \(W\) is a constant matrix with \(W_{i,j} = 1\), so that \(B\) is block-constant. One eigenvector of \(W\) is \(w^{(1)} = [1,1,\cdots,1]^T\) with eigenvalue \(\omega_1 = P\). A choice for the other \(P - 1\) linearly-independent eigenvectors is \(w^{(2)} = [1,-1,0,\cdots,1]^T\), \(w^{(3)} = [1,0,-1,0,\cdots,1]^T\), etc., all with identically-zero eigenvalue. Equation (2) implies that the eigenvalues of \(B\) consist of the \(M\) eigenvalues of \(A\), all multiplied by \(P\), together with \(M - P\) identically-zero eigenvalues. Equation (2) also implies that the corresponding eigenvectors of \(B\) consist of “interpolated” versions of the \(M\) eigenvectors of \(A\) (with constant interpolation), together with eigenvectors \(b_j^{(s)}\) (with \(M < s \leq N\)) that oscillate rapidly with \(j\).

3. Interpolation of a Timeseries. Suppose that the time series \(m\) has sampling interval \(\Delta t\) and length \(M\). Suppose also that the time series \(y\) has sampling interval \(\Delta t/P\) and length \(N = PM\). Consider a \(N \times M\) interpolation operator \(T\) that takes \(m\) into \(y\) (i.e. \(y = Tm\)) and that preserves the values of \(m\). That is, \(T_{ip,j} = \delta_{i,j}\) so that \(y_{kp} = m_k\), for \(0 \leq k < M\). The \(P = 2\) case corresponding to linear interpolation is:

\[
T = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & \cdots \\
0 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{bmatrix}
\]

(5)

A symmetric \(N \times N\) matrix \(B\) can be formed from a symmetric \(M \times M\) matrix \(A\) via:

\[
B = TAT^T
\]

(6)

This matrix preserves the values of \(A\) in \(B\):

\[
B_{kp,kp} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} T_{kp,i}T_{kp,j}A_{i,j} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \delta_{k,i}\delta_{k,j}A_{i,j} = A_{i,j}
\]

(7)
and interpolates the other values. Furthermore, the interpolation process can be viewed as acting on the eigenvectors of $\mathbf{A}$, since it follows from $\mathbf{A} = \sum_s \alpha_s \mathbf{a}^{(s)}\mathbf{a}^{(s)\mathsf{T}}$ that:

$$
\mathbf{B} = \mathbf{T}\mathbf{A}\mathbf{T}^\mathsf{T} = \sum_{s=0}^{M-1} \alpha_s \mathbf{a}^{(s)}\mathbf{a}^{(s)\mathsf{T}}\mathbf{T}^\mathsf{T} = \sum_{s=0}^{M-1} \alpha_s (\mathbf{a}^{(s)})(\mathbf{a}^{(s)})^\mathsf{T}
$$

Note, however that $\mathbf{T}\mathbf{a}^{(s)}$ is not, in general, an eigenvector of $\mathbf{B}$.

Now consider forming the covariance $\mathbf{C}_y$ of $\mathbf{y}$ from the covariance matrix $\mathbf{C}_m$ of $\mathbf{m}$ using the usual rules of error propagation:

$$
\mathbf{C}_y = \mathbf{T}\mathbf{C}_m\mathbf{T}^\mathsf{T}
$$

Comparing Equations (6) and (8), we conclude that $\mathbf{C}_y$ is an interpolated version of $\mathbf{C}_m$. We expect that $\mathbf{C}_y$ has at least $N - M$ identically-zero eigenvalues, because the interpolation has created $N - M$ linear elements $\mathbf{y}$ that are completely dependent on the other $M$ elements and that, consequently, have no error.

4. Effect of Interpolation on Eigenvalues and Eigenvectors. We now address the question of how the eigenvalues and eigenvectors of $\mathbf{C}_y$ differ from those of $\mathbf{C}_m$. Since we have established that $\mathbf{C}_y$ is an interpolated version of $\mathbf{C}_m$, we start by writing:

$$
\mathbf{C}_y = \mathbf{C}_y^{(0)} + \delta\mathbf{C}_y
$$

We choose $\mathbf{C}_y^{(0)}$ to be a block-constant matrix built from $\mathbf{C}_m$ (with block size $P$). We have previously established that this block-constant matrix shares the eigenvectors and (up to a multiplicative constant) eigenvalues of $\mathbf{C}_m$, and additionally has $N - M$ highly oscillatory eigenvectors with identically-zero eigenvalues. The perturbation $\delta\mathbf{C}_y$ does not necessarily have identically-zero mean, either overall or within individual $P \times P$ blocks. However, since $\mathbf{C}_y$ shares every $P$th value with $\mathbf{C}_y^{(0)}$ (the others being determined by interpolation), we expect that these means are very much smaller than $\|\delta\mathbf{C}_y\|$.

The effect of the small perturbation $\delta\mathbf{C}_y$ on the non-zero eigenvalues can be determined using perturbation theory. The eigenvalues and eigenvectors of $\mathbf{C}_m$ are denoted as $\alpha_s$ and $\mathbf{a}^{(s)}$, respectively, and the eigenvalues and eigenvectors of $\mathbf{C}_y$ as $\beta_s = \beta_s^{(0)} + \delta\beta_s$ and $\mathbf{b}^{(s)} = \mathbf{b}^{(0s)} + \delta\mathbf{b}^{(s)}$, respectively. The $M$ non-zero eigenvalues are non-degenerate, so the first-order perturbations can be shown to be:

$$
\delta\beta_s = \mathbf{b}^{(0s)\mathsf{T}}\delta\mathbf{C}_y \mathbf{b}^{(0s)} \quad \text{and} \quad \delta\mathbf{b}^{(s)} = \sum_{j=0}^{N-1} \left( \frac{\mathbf{b}^{(0s)\mathsf{T}}\delta\mathbf{C}_y \mathbf{b}^{(0j)}}{\beta_s^{(0)}} \right) \mathbf{b}^{(0j)}
$$
We consider these first-order perturbations to be “small”. In fact, $\delta \beta_s$ is especially small: Because $b^{(0s)}$ is block-constant, the $\delta \beta_s$ would be zero if each block of $\delta C_y$ has zero mean:

$$
\delta \beta_s = b^{(0s)T} \delta C_y b^{(0s)} = \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} b_p^{(0s)} [\delta C_y]_{p,q} b_q^{(0s)}
$$

$$
= \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} a_i^{(s/P)} a_j^{(s/P)} w_u^{(s/P)} [\delta C_y]_{i,j} a_i^{(s/P)} a_j^{(s/P)} w_u^{(s/P)}
$$

$$
= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_i^{(s/P)} a_j^{(s/P)} \sum_{u=0}^{P-1} \sum_{v=0}^{P-1} w_u^{(s/P)} [\delta C_y]_{i,j,u,v} w_u^{(s/P)}
$$

$$
= p^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_i^{(s/P)} a_j^{(s/P)} \sum_{u=0}^{P-1} \sum_{v=0}^{P-1} [\delta C_y]_{i,j,u,v} = 0
$$

when $\sum_{u=0}^{P-1} \sum_{v=0}^{P-1} [\delta C_y]_{i,j,u,v} = 0$ for all $i,j$

The term in parenthesis is the block-mean. Here we have used the fact that, for the non-zero eigenvalues, $w_v^{(s/P)} = w_v^{(s/P)} = P$. In actuality, the block-means are almost but not quite zero, so that $\delta \beta_s$ is non-zero but much smaller than $\|b^{(0s)}\|^2 \|\delta C_y\|$.

The $N - M$ zero eigenvalues are degenerate. Their perturbations also be calculated using degenerate perturbation theory, but we omit discussion of it here.

5. Effect on Menke & Blatter Style Trade-Off Curves. The eigenvalue spectrum of $C_y$ differs from that of $C_m$ in three ways: an overall scaling factor of $P$ is introduce that represents the decrease of the sampling interval from $\Delta t$ to $\Delta t / P$; the $M$ non-zero eigenvalues of $C_y$ are slightly perturbed with respect to those of $C_m$; and $N - M$ zero eigenvalues are added. Superficially, the introduction of the zero eigenvalues might seem to be a boon, since they imply that zero-variance features have been added to the problem. However, these features arise from the interpolation and are associated with highly-oscillatory eigenvectors. Thus, given $[y_1, y_2, y_3, y_4, y_5, \cdots]^T$ with the linear interpolation $y_2 = \frac{1}{2}y_1 + \frac{1}{2}y_3$ and $y_4 = \frac{1}{2}y_3 + \frac{1}{2}y_5$, the linear combinations $\frac{1}{2}y_1 - y_2 + \frac{1}{2}y_3 = 0$ and $\frac{1}{2}y_3 - y_4 + \frac{1}{2}y_5 = 0$ have identically-zero variance. However, they are also oscillatory and highly unlocalized and cannot be used to form a localized weighted average. For instance, although the sum of these two linear combinations, say $\langle y_3 \rangle = \frac{1}{2}y_1 - y_2 + y_3 - y_4 + \frac{1}{2}y_5$ is centered about $y_3$, it is not usefully localized around $y_3$. Consequently, the $N - M$ zero eigenvalues merely add a long tail of small-variance, large-spread values to the trade-off curve. The part of the trade-off curve with small-spread is controlled by the $M$ non-zero eigenvalues, and since these are only slightly perturbed with
with respect to those of $\mathbf{c}_m$, this part of the trade-off curve for $\mathbf{y}$ is very similar to that for $\mathbf{m}$ (up to an overall scaling).

The upshot is that a sound interpretation of variance and resolution can be made from $\mathbf{y}$. 